DOCUMENT RESUME

ED 359 940 IR 016 160

TITLE	High Performance Computing: Advanced Research Projects Agency Should Do More To Foster Program Goals. Report to the Chairman, Committee on Armed Services, House of Representatives.
INSTITUTION	General Accounting Office, Washington, DC. Information Management and Technology Div.
REPORT NO	GAO/IMTEC-93-24
PUB DATE	May 93
NOTE	45p.
AVAILABLE FROM	U.S. General Accounting Office, P.O. Box 6015, Gaithersburg, MD 20884-6015 (first copy free; $2 for each additional copy; 100 or more to a single address discounted 25 percent).
PUB TYPE	Reports - Descriptive (141) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Agency Role; *Computer Oriented Programs; Computer Software; Computer Software Development; Computer Software Selection; *Federal Programs; Guidelines; Information Dissemination; Interviews; *Research and Development; Research Methodology; Research Projects; *Technological Advancement
IDENTIFIERS	Advanced Research Projects Agency; General Accounting Office; *High Performance Computing; High Performance Computing Act 1991; *Supercomputers

ABSTRACT

High-performance computing refers to the use of advanced computing technologies to solve highly complex problems in the shortest possible time. The federal High Performance Computing and Communications Initiative of the Advanced Research Project Agency (ARPA) attempts to accelerate availability and use of high performance computers and networks. The House Armed Services Committee asked the General Accounting Office (GAO) to assess this program, particularly the agency's distribution of advanced computers to research sites, its interaction with the research community, and the balance between hardware and software. Interviews with agency and industry representatives and a review of documents were conducted. Results indicate that the ARPA has fostered significant advances, but with several shortcomings. Placing of new computers has focused on just a few machines. Limited interactions with the technical community may be hindering progress toward ambitious program goals. Much progress has been made in hardware development, but software remains too primitive to make massively parallel processing systems useful. Recommendations for improvement are presented. Four appendixes provide details about methodology, placement of machines, and project selection, and list contributors to the report. Five tables contain study findings, and two figures illustrate computer applications and product selection. (SLD)

United States General Accounting Office

GAO

Report to the Chairman, Committee on Armed Services, House of Representatives

May 1993

HIGH PERFORMANCE COMPUTING

Advanced Research Projects Agency Should Do More to Foster Program Goals

ED359940

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

GAO/IMTEC-93-24

GAO

United States
General Accounting Office
Washington, D.C. 20548

Information Management and
Technology Division

B-252539

May 17, 1993

The Honorable Ronald V. Dellums
Chairman, Committee on Armed Services
House of Representatives

Dear Mr. Chairman:

This report responds to your request that we review the investment strategy of the Advanced Research Projects Agency's (ARPA) High Performance Computing (HPC) Program. Specifically, you asked that we assess (1) ARPA's distribution of advanced computers to research sites, (2) ARPA's interaction with the research community, and (3) the balance between software and hardware investments in the ARPA program. The report identifies weaknesses in the program, which could slow technological progress and prevent ARPA from achieving its HPC program goals, and makes recommendations to the Secretary of Defense to improve the program.

We will give copies of the report to appropriate congressional committees, the Secretary of Defense, and other interested parties. Copies will also be made available to others upon request.

This work was performed under the direction of Samuel W. Bowlin, Director, Defense and Security Information Systems, who can be reached at (202) 512-6240. Other major contributors are listed in appendix IV.

Sincerely yours,

Ralph V. Carlone

Ralph V. Carlone
Assistant Comptroller General

Executive Summary

Purpose

High-performance computing refers to the use of advanced computing technologies, especially supercomputers, to solve highly complex, numerically intensive problems in the shortest possible time. These scientific problems—such as understanding global climate change or analyzing molecular structure—are collectively called the grand challenges. The federal High Performance Computing and Communications Initiative is a research and development effort that seeks to significantly accelerate the availability and utilization of high performance computers and networks in order to better address these challenges. At $275 million in fiscal year 1993, the Advanced Research Projects Agency (ARPA) has the largest budget of any single agency participating in the $800-million federal initiative.[1]

Given the importance of this initiative and ARPA's dominant role in it, the House Armed Services Committee asked GAO to assess the program, particularly (1) the agency's distribution of advanced computers to research sites, (2) its interaction with the research community, and (3) the balance between hardware and software investments in the ARPA program.

Background

ARPA has been funding high-performance computing research and development since it began a strategic computing program in the early 1980s, and is now one of the lead agencies in the federal initiative. It is the primary agency involved in the research and development of critical high-performance computing technology, needed to address the grand challenges. ARPA coordinates its projects with other agencies, especially the National Science Foundation, which focus more on applications, infrastructure, and education. ARPA funds some 200 projects in this area of inquiry, half managed by industry and half by universities.

ARPA has concentrated on a new approach to supercomputer design, based on the interconnection of hundreds or even thousands of microprocessors; this is commonly known as massively parallel processing. Although ARPA's achievements in computing technology are widely recognized, its program in this specific area has been controversial because of its seemingly narrow emphasis on increasing machine speed and because of its continuing support for only a few select vendors. The high-performance computing program is also unusual for the agency in that its goals go beyond basic research, stressing development of useful massively parallel processing systems.

[1]Until March 1993, ARPA was known as the Defense Advanced Research Projects Agency.

Results in Brief

While ARPA has fostered significant advances in high-performance computing research and development for a decade, its current program has shortcomings in several areas. First, its placing of new computers at laboratories, while important to facilitating research, has focused on just a few massively parallel processing machines. Researchers need access to a broader range of new computing technologies to explore all promising alternatives. Second, the agency's limited interactions with the technical community may hinder the rapid progress needed to achieve its ambitious program goals. ARPA has been weak in disseminating program information, soliciting input from the research community, and publishing performance data. Finally, while much progress has been made in hardware development, software remains too primitive to make massively parallel processing systems useful. ARPA's goal of achieving a thousand-fold increase in useful computing power by 1996 will likely not be met without greater emphasis on the development of system software.[2]

Principal Findings

ARPA Has Been Placing Too Narrow a Range of Computers

ARPA has actively sought to place new massively parallel processing machines that it has helped to develop into the hands of researchers as quickly as possible. This practice of placing early prototype machines in research settings is widely supported in theory, yet the agency's specific actions have been heavily criticized as biased toward the products of two vendors who have received research and development funding from the agency.

To date, ARPA's high-performance computing and strategic computing programs have facilitated the procurement of computers made by its development contractors, including 44 systems made by Intel Corporation, and 24 systems made by Thinking Machines Corporation, as well as a number of computers made by companies that are no longer marketing massively parallel processing products. However, ARPA has not been involved in any major procurement of new machines made by current major rivals to Intel and Thinking Machines. Since it is important that a broad variety of massively parallel processing designs be made available to

[2]System software is the collection of programs and data that make up and relate to the operating system, for example, input/output routines, command-line interpreters, and task scheduling and memory management routines. Application software is software designed to fulfill the specific needs of a user.

researchers so that all promising approaches can be tested, ARPA's placing of only the products of its development contractors is not justified.

Limited Interaction With Research Community Hinders Progress

ARPA's relatively large high-performance computing budget and past computing accomplishments make it a leader in massively parallel processing research and development. Within the federal initiative, the agency's role is also dominant; other agencies look to ARPA-sponsored research and development for technological advances that will support their own efforts. Rapid progress in high-performance computing hinges on ARPA's effectively interacting with the broad spectrum of researchers from government laboratories, academia, and industry. Such interactions with this community have, however, been limited.

ARPA gets input from the research community chiefly by interacting with its own principal investigators. It publishes no detailed summaries or progress reports. Researchers likewise find it difficult to understand how ARPA selects projects for funding. At the other end of the process, individual agency-sponsored projects are not required to publish results in any standard format or at any predetermined time. The lack of widely available performance data on new designs slows technological progress because the research community remains uncertain about the merits of the new designs.

ARPA Has Not Sufficiently Addressed System Software

The development of system software that would enable researchers to make full use of the tremendous processing power of massively parallel processing machines has not kept pace with the development of hardware. This shortcoming threatens attainment of ARPA's goal of achieving a thousand-fold increase in useful computing power by 1996.

High-performance computing experts from academia, government, and industry have criticized the agency for overemphasizing hardware. Specific system software areas needing greater attention have been identified, including programming languages, compilers, and program development tools.[3] According to these experts, ARPA must focus more on system software development to meet its program goals.

[3]Compilers are system software programs that translate source code, written in high-level programming languages, into machine-executable object code.

Recommendations

In order to broaden participation in ARPA's program and to better facilitate research into high-performance computing, GAO recommends that the Secretary of Defense direct the Director, ARPA to (1) broaden the agency's computer placement program by including a wider range of computers from more vendors; (2) establish and maintain a public database of information about the status and results of the agency's ongoing high-performance computing projects, as well as performance data for different massively parallel processing systems; and (3) emphasize and support research and development of system software as a major element of the agency's high-performance computing program. Other related recommendations are included in chapters 2 through 4 of the report.

Agency Comments

GAO did not obtain written agency comments on a draft of this report. However, GAO provided a draft of the report to ARPA and discussed the report's findings and recommendations with ARPA officials, including the Director, Deputy Director for Management, and High Performance Computing Program Director. Their specific comments have been incorporated in the text and revisions to the report have been made as appropriate. The ARPA officials generally agreed that improvements could be made in the management of the high-performance computing program. Regarding placement of computer systems at research sites, the ARPA officials noted that they are already taking some steps to broaden their program. With respect to GAO's finding that ARPA's interaction with researchers has been limited, the ARPA officials stated that their dissemination of public information is better than the report indicates. Regarding GAO's finding that the agency has not sufficiently addressed system software, ARPA officials stated that they have invested substantially in software development. In each of these cases, GAO believes that ARPA's actions have not been sufficient to invalidate the findings of the report. The steps ARPA is taking to broaden its computer placement program are still in the early stages; GAO believes attention to this issue is still needed. ARPA's efforts to disseminate public information still lack a mechanism for providing detailed performance results from technology development projects. Provisions for obtaining broad scientific input to the program are also still lacking. Finally, notwithstanding ARPA's investment to date in software, GAO found during this review that a substantially greater emphasis needs to be placed on system software development if ARPA's program goals are to be met. ARPA officials' oral comments and GAO's evaluation of them are included at the end of each chapter.

Contents

Executive Summary		2
Chapter 1 Introduction		8
	The Governmentwide Initiative	8
	ARPA's HPC Program	12
	Massively Parallel Processing	13
	ARPA's Goals Are Ambitious	14
	The MPP Research and Development Community	14
Chapter 2 ARPA's Approach to Distribution of MPP Prototype Machines May Limit Options for Government Researchers		16
	The Need for Rapid Deployment of New Computers	16
	ARPA's Program	17
	Industry and Researchers Have Questioned ARPA's Placement of MPP Machines	17
	Conclusions	19
	Recommendations	20
	Agency Comments	20
Chapter 3 ARPA's HPC Program Would Benefit From Greater Collaboration With the HPCC Research Community		21
	ARPA's Interaction With the Research Community Is Too Limited	21
	Project Results Are Not Published in a Uniform and Timely Fashion	22
	MPP Performance Measurement Standards Are Lacking	23
	Conclusions	24
	Recommendations	24
	Agency Comments	24
Chapter 4 ARPA Must Emphasize Software to Meet Its Goals		26
	Experts Have Emphasized the Need for Greater Investment in System Software	26
	Primitive System Software Limits MPP's Usefulness	27
	ARPA's Approach Focuses on Early Phases of Software Development	29
	Conclusions	29
	Recommendations	30
	Agency Comments	30

Appendixes

Appendix I: Objectives, Scope, and Methodology 32
Appendix II: ARPA's Placement of Intel and Thinking Machines Products 34
Appendix III: ARPA's HPC Project Selection Process 37
Appendix IV: Major Contributors to This Report 42

Tables

Table 1.1: HPCC Initiative Fiscal Year 1993 Budget by Agency and Component 9
Table II.1: Intel Touchstone Machines Placed by ARPA 34
Table II.2: Intel iWarp Machines Placed by ARPA 35
Table II.3: Thinking Machines Corp. Connection Machines Placed by ARPA 36
Table III.1: Proposals Awarded Under Synopsis #57 41

Figures

Figure 1.1: Examples of High Performance Computing Applications 10
Figure III.1: ARPA's HPC Project Selection Process 38

Abbreviations

ARPA	Advanced Research Projects Agency
ASTA	Advanced Software Technology and Algorithms
BAA	broad agency announcement
BRHR	Basic Research and Human Resources
DOD	Department of Defense
DOE	Department of Energy
EPA	Environmental Protection Agency
GAO	General Accounting Office
HPC	high performance computing
HPCC	High Performance Computing and Communications
HPCS	High Performance Computing Systems
IMTEC	Information Management and Technology Division
MPP	massively parallel processing
NASA	National Aeronautics and Space Administration
NIH	National Institutes of Health
NIST	National Institute of Standards and Technology
NOAA	National Oceanic and Atmospheric Administration
NREN	National Research and Education Network
NSA	National Security Agency
NSF	National Science Foundation

Introduction

The term high-performance computing (HPC) refers to the use of advanced computing, communications, and information technologies, such as supercomputers and high-speed networks, to solve highly complex, mathematically intensive problems in the shortest possible time. HPC is widely considered to be a critical technology area that can further progress in solving fundamental problems that are critical to ensuring competitiveness and that require significant increases in computational capability. Examples include prediction of global climate change, determination of molecular structure, and understanding the nature of new synthetic materials. Some of these examples are shown in figure 1.1. Because HPC is so important, the federal government plays a leading role in its development.

The Governmentwide Initiative

The federal High Performance Computing and Communications (HPCC) Initiative began in fiscal year 1992 as a joint effort among nine federal agencies to significantly accelerate the availability and utilization of the next generation of high-performance computers and networks. Specifically, the initiative aims for a thousand-fold improvement in useful computing capability and a hundred-fold improvement in available computer communications capability by 1996. There are four major components of the HPCC initiative:

- High Performance Computing Systems (HPCS): the development of the underlying technology required to build systems capable of sustaining trillions of operations per second on large problems;
- Advanced Software Technology and Algorithms (ASTA): the development of software technology and algorithms to address grand challenge problems;
- National Research and Education Network (NREN): the development of a national high-speed network to provide distributed computing capability to research and educational institutions; and
- Basic Research and Human Resources (BRHR): support for research by individual investigators and initiation of activities to increase the pool of trained personnel.

Planned fiscal year 1993 spending by agency and HPCC component is shown in table 1.1.

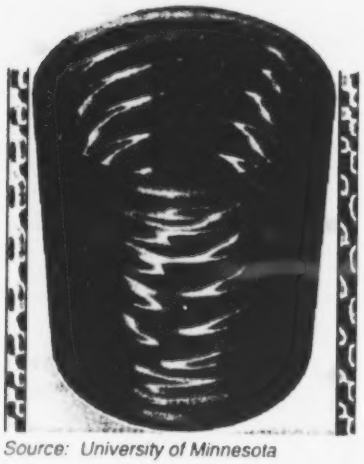
Table 1.1: HPCC Initiative Fiscal Year 1993 Budget by Agency and Component

Agency	HPCS	ASTA	NREN	BRHR	TOTAL
ARPA	119.5	49.7	43.6	62.2	275.0
NSF	25.8	107.6	41.1	50.6	225.1
DOE	10.1	63.8	13.8	12.9	100.6
NASA	12.0	57.6	9.1	3.1	81.8
NSA[a]	36.2	5.9	3.2	0.2	45.5
NIH	3.0	31.4	4.1	8.0	46.5
NOAA	0.0	9.4	0.4	0.0	9.8
EPA	0.0	6.0	0.4	1.5	7.9
Education	0.0	0.0	2.0	0.0	2.0
NIST	0.3	0.6	1.2	0.0	2.1
Total	**206.9**	**332.0**	**118.9**	**138.5**	**796.3**

[a]NSA joined the HPCC Initiative in 1993. Previously it had been represented by ARPA.

No single federal agency has overall responsibility for the initiative. Instead, participating agencies coordinate their program plans and activities through an arm of the White House Office of Science and Technology Policy, which establishes task groups, prepares budgets, and coordinates program planning and execution. The four agencies involved in preparing the original HPCC plan—the Department of Energy (DOE), the Advanced Research Projects Agency (ARPA), the National Aeronautics and Space Administration (NASA), and the National Science Foundation (NSF)—are the major participants in the initiative.

Figure 1.1: Examples of High Performance Computing Applications



Source: University of Minnesota

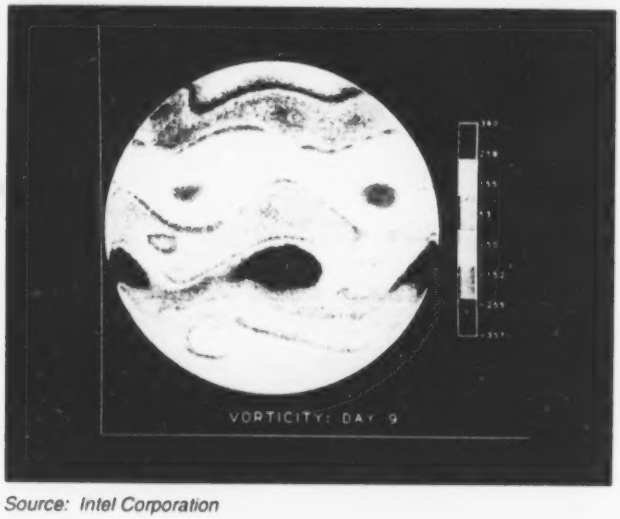
A Thinking Machines CM-5 was used for this study of fluid flow between two concentric cylinders.

Source: Intel Corporation

An Intel iPSC/860 was used by researchers at Argonne National Laboratory to study the molecular structure of potential enzyme designs.

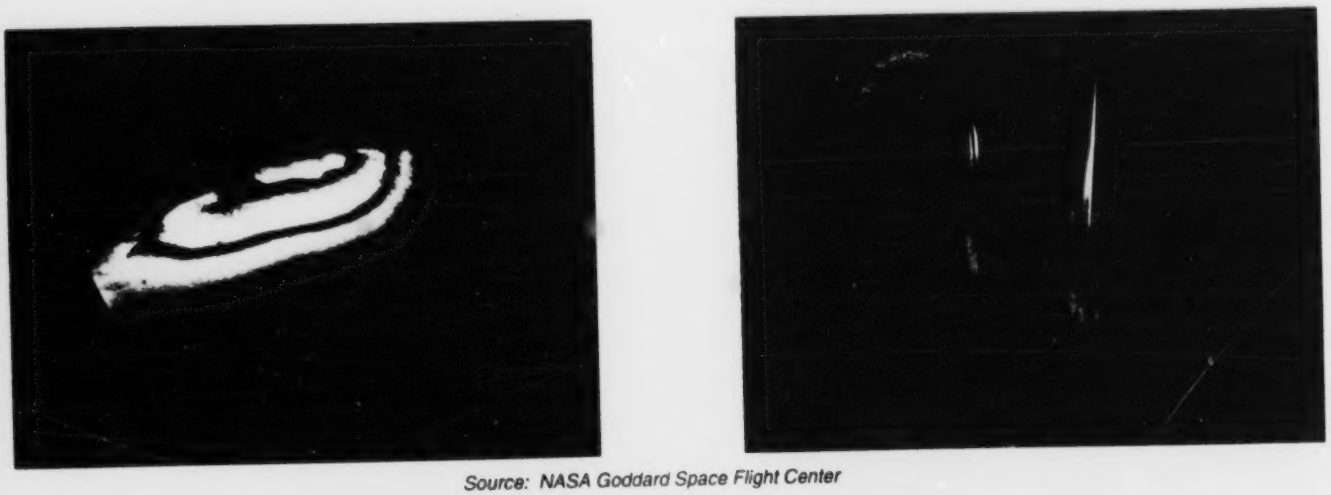
Source: nCUBE Corporation

Researchers at Sandia National Laboratories used an nCUBE MPP to model airflow as the space shuttle reenters the Earth's atmosphere at 8 times the speed of sound.



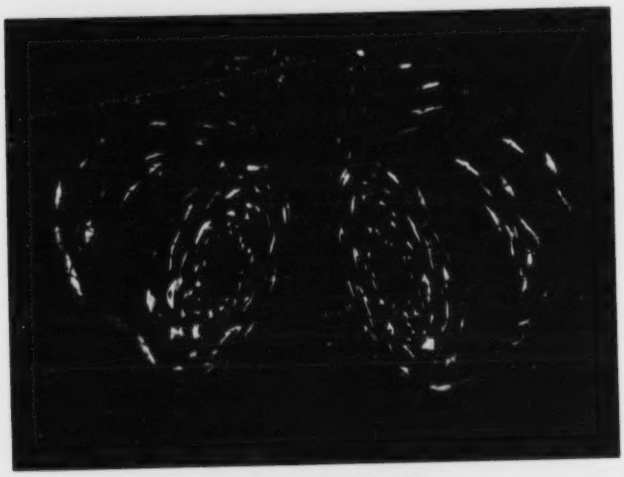
Source: Intel Corporation

Researchers used the Intel Delta at the California Institute of Technology for this new approach to global climate modeling.



Source: NASA Goddard Space Flight Center

Use of a MasPar MPP allowed scientists to discern a binary star formation (on right) that was hidden in the original (on left) taken by the Hubble Space Telescope.



Source: Lawrence Livermore National Laboratory

This study of density fluctuations in a gaseous structure used for nuclear fusion was performed on a Thinking Machines CM-2.

BEST COPY AVAILABLE

ARPA's HPC Program

ARPA is the lead agency within the Department of Defense (DOD) for advanced technology research, and ARPA represents DOD's interests within the federal HPCC initiative. In general, ARPA is tasked with pursuing imaginative and innovative research and development projects that promise significant military utility. ARPA's purview includes basic and applied research and development projects that demonstrate the feasibility of revolutionary technological approaches.

ARPA has a long history of championing key developments in the evolution of computer technology. For example, the agency was instrumental in developing a computer communications networking approach known as packet switching during the 1960s and 1970s. When ARPA first began building a packet-switched computer network, the technique was unproven; established commercial firms were not interested in financing packet-switching research and development on their own. However, due in large part to the success of ARPA's initiative, the technique eventually came into widespread use and has facilitated the interconnection of heterogeneous computer systems and resources around the world. ARPA has also been involved in the development of other important technologies, such as advanced interactive computer graphics and various aspects of artificial intelligence.

ARPA's Computing Systems Technology Office administers the agency's HPC program. The program consists of a number of research and development projects that are carried out by both academic researchers and commercial developers. Projects range from small-scale experiments involving new hardware and software design concepts to large-scale development of complete computer systems. Currently, ARPA funds some 200 projects in HPC, which are split fairly evenly between industry and university-sponsored projects. At $275 million in fiscal year 1993, ARPA controls the single largest share of HPCC resources.

Officials from various agencies involved in the HPCC initiative agree that ARPA is the de facto leader in supporting research and development of new HPC technology. The agency has been supporting HPC research and development since the early 1980s, when it initiated the Strategic Computing Program. One of the Strategic Computing Program's original goals was to expand practical experience with experimental parallel computer systems. In the 1980s, ARPA's support was critical to the development of several massively parallel processing (MPP) computers, including the Connection Machine and the Intel Touchstone series. Prior to this, parallel processing had been largely ignored by the scientific

research community, and few manufacturers had offered MPP products. ARPA has maintained a steady investment in MPP research since that time and has largely succeeded in proving the overall concept and feasibility of the MPP approach. The early success of ARPA's parallel computing research led ARPA officials to become heavily involved in the formulation of the federal HPCC initiative.

Massively Parallel Processing

MPP represents a revolutionary new computer design. Most traditional computers have one computational processor, and traditional computer development has focused on making this processor faster and more efficient. However, the potential for continued increases in speed is reaching the limits imposed by the physical properties of the materials used to build the processor.

Most supercomputers currently used to solve "real-world" problems are vector supercomputers. Vector machines achieve much greater speeds than traditional computers by employing processors that process data in an assembly-line fashion. In a vector processing machine, a number of instructions are in various stages of being processed at any given moment as they make their way through the machine's high-speed processing pipelines. Results emerge at the end of these pipelines just as completed automobiles emerge from automotive assembly lines. Computers manufactured by Cray Research, Inc., are typical of this kind of supercomputer. While these machines are much more powerful than traditional single-processor, general-purpose computers, they are still not powerful enough to adequately address grand challenge problems.

Many computer scientists believe that MPP promises speeds far surpassing those of vector supercomputers by breaking computational problems into many separate parts and having a large number of processors tackle those parts simultaneously. Speed is achieved largely through the sheer number of processors operating simultaneously, rather than through any exceptional power in each processor. In fact, many MPP designs use commercial, off-the-shelf processors, such as those found in personal computers or scientific workstations, and may include hundreds or even thousands of these processors.

Most MPP designs are intended to be scalable; that is, the machines function effectively in configurations that range from a small number of processors to a very large number of processors. While the number of processors may vary, the system's basic architecture and system software

are constant. Thus these machines can be tailored to match a wide variety of computing demands.

The concept of massive parallelism can be implemented in many different ways. Efficient methods must be developed to break up large computational problems and assign tasks to individual processors. Likewise, new methods must be devised for efficiently managing communications among the processors. Although some solutions to these problems seem clearly better than others, there is still no consensus about what approaches are best overall. Getting large amounts of data quickly and efficiently sent to and retrieved from a large collection of processors and their associated memory modules also presents significant engineering challenges.

ARPA's Goals Are Ambitious

The goal of ARPA's HPC program is to develop systems capable of sustaining trillions of operations per second on grand challenge problems by 1996. Current vector supercomputers offer billions of operations per second of useful computing power; the ARPA-sponsored systems would thus represent a thousand-fold improvement in computational capability. ARPA's objective is not only to achieve very high processing speeds in laboratory tests, but to make that computing power useful in addressing complex real-world problems.

ARPA's goals are ambitious because its officials envision a substantial conversion of scientific research from vector to massively parallel supercomputing. However, much of the theoretical improvement in speed and processing efficiency expected of MPP machines has yet to be realized. Current MPP machines have demonstrated superiority to vector supercomputers on only a few applications. In order for MPP to gain widespread acceptance in the scientific research community, its effectiveness on large, complex scientific problems will have to be demonstrated.

The MPP Research and Development Community

Typically, MPP computers have been manufactured by small companies or corporate divisions established specifically to research, develop, and produce a particular MPP design. Examples include Intel Supercomputer Systems Division, Kendall Square Research Corporation, MasPar Computer Corporation, Meiko Scientific Corporation, nCUBE Corporation, Parsytec GmbH, and Thinking Machines Corporation. Because the market for these computers is very limited, the few

companies that are actively marketing products compete vigorously for each sale. Larger, more established computer manufacturers have been slower to enter the MPP field but are becoming more interested as the technology matures. Examples of large computer manufacturers that are now actively involved in MPP include Cray Research, Incorporated, Digital Equipment Corporation, and International Business Machines, Incorporated.

Although some MPP machines are in operation, achieving such goals as a thousand-fold increase in useful computing is still a subject of research and development. Within the government, MPP research and development are conducted at DOE and DOD laboratories, at NASA centers, and at NSF supercomputer centers. Typically, these government sites have extensive experience with traditional vector supercomputers and support large numbers of users. Because these sites will eventually shift their users to MPP machines, many of them operate small laboratories where researchers evaluate how well new MPP machines handle problems that typify their interests.

ARPA's Approach to Distribution of MPP Prototype Machines May Limit Options for Government Researchers

ARPA has been actively involved in providing researchers with quick access to early prototypes of MPP machines that it has helped to develop. By testing the new machines on real-world problems, researchers can rapidly identify the machine's strengths and weaknesses and provide feedback to system developers. While this activity is widely supported in concept, ARPA's specific actions have been criticized as too narrowly focused on the products of vendors who have received research and development funding from ARPA. To date, ARPA has facilitated the acquisition of a relatively large number of computers made by Intel Supercomputing Systems Division and Thinking Machines Corporation, but has not provided to researchers machines made by major rivals to these companies, including Kendall Square Research Corporation, nCUBE Corporation, and MasPar Computer Corporation. Some researchers and vendors have charged that ARPA's wide placement of Intel and Thinking Machines products has made it difficult for researchers to gain access to rival machines. By making a broader selection of alternative MPP designs available, ARPA would not only dispel this controversy but also better ensure that promising MPP design alternatives are not overlooked.

The Need for Rapid Deployment of New Computers

Research and development in the field of massively parallel processing is progressing at a rapid pace. Members of the MPP community, including government agencies, laboratories, and vendors, vigorously compete with each other to be at the forefront of the field. In order for rapid progress to occur, new MPP machines must not only be built but must also be made accessible to researchers and thoroughly tested so that their strengths and weaknesses can be identified. The results of this experimental testing can then be used to influence future versions of the machines.

If new machines are built but not provided to the community, real-world experiments do not occur and significant technical progress cannot be made. In such cases, the government has funded a laboratory curiosity—a single prototype that few researchers can access and that contributes little to the body of engineering design experience in the field.

Researchers generally view the government's standard acquisition process for computer equipment as a barrier to effective participation in HPC research. According to these researchers, standard government procurement procedures for acquiring major computer systems can take as many as 2 to 3 years to complete, by which time a new generation of MPP machines will have been produced. If no alternative acquisition process were available, government laboratories could not effectively

participate in HPC research. Because some agencies have no special procedures for such time-critical, research-oriented procurements, ARPA assists them, using special procedures for rapidly acquiring advanced computer prototypes.

ARPA's Program

ARPA works closely with industry and academia to develop new technologies, such as MPP. However, it does not operate any laboratories of its own where these technologies can be tested. Instead, ARPA forms partnerships with other government agencies, laboratories, and industry to evaluate the technology it helps to develop. ARPA places computer systems at sites controlled by these other organizations based on proposals it receives from them or as part of joint projects with other government agencies. Systems placed at government and academic sites may be paid for by either ARPA or the site, and ARPA's involvement may bring a lower price, since ARPA has ongoing contractual relationships with some MPP developers. The site is often responsible for operations and maintenance expenses.

Almost all of the MPP systems that have been placed so far by ARPA were manufactured by either Intel or Thinking Machines.[1] To date, ARPA's HPC and Strategic Computing programs have facilitated the procurement of 44 computer systems made by Intel and 24 systems made by Thinking Machines. Appendix II contains complete lists of these placements, including the sites where the machines were placed, their cost, and the source of the funding. ARPA supported the development of both vendors' machines. As part of the development contracts for these machines, ARPA negotiated options that would allow the government to buy a certain number of "early prototype copies" at a discount for evaluation purposes.

Industry and Researchers Have Questioned ARPA's Placement of MPP Machines

The fact that ARPA is able to provide machines quickly to the research community is something that many researchers value. Nevertheless, much controversy surrounds ARPA's placement of particular MPP machines. A number of government officials, researchers, and vendors have said that ARPA's actions have distorted the market for MPP machines because ARPA has almost exclusively favored the products of just two companies, Intel and Thinking Machines. The ARPA HPC program has not supported any procurements of machines made by several other major MPP developers—such as Kendall Square Research Corporation, nCUBE

[1]In the late 1980s, ARPA also placed machines made by vendors who have since left the MPP field, including BBN Advanced Computers, Inc. and Encore Computer Company.

Corporation, and MasPar Computer Corporation—with whom ARPA did not have MPP development contracts.

Industry representatives are concerned that the Intel and Thinking Machines products placed by ARPA largely fall into the category of platforms for research into the applied use of MPP instead of early prototypes to be rigorously tested for engineering soundness. Researchers and industry representatives whom we interviewed generally agreed that the placement of just two or three machines at key research and development sites would be sufficient for rapid testing of early prototypes. ARPA has placed many more Touchstones, iWarps, and Connection Machines than this. Researchers argue that all eligible MPP machines should be involved in the placement program if the emphasis is on applied research rather than early prototype testing.

While some of the placements of the Intel and Thinking Machines products were clearly intended for early testing of prototype versions, other placements have been used largely for research into the applicability of parallel processing to various scientific disciplines. For example, two of the Intel machines, a Touchstone and an iWarp, were placed by ARPA at the Naval Ocean Systems Center's Center for Advanced Computation. Navy documents indicate that the primary purpose of the center is to explore new MPP applications, such as anti-submarine warfare, not to advance computer science per se. Other sites also fit the same profile.

Design and engineering information about the Intel and Thinking Machines products also indicates that later editions of the machines acquired through ARPA were significantly more refined than the first few machines produced by each company. Intel's first two deliveries under the Touchstone program, to NASA Ames Research Center and Oak Ridge National Laboratory, were clearly early prototypes intended for field testing. The machines were marked "Touchstone Gamma," the official designation of the prototype program, and, according to Intel officials, were significantly less refined from an engineering standpoint than later copies of the machines. The additional Touchstone machines placed by ARPA were all marked "iPSC/860"—Intel's commercial designation for the machine—and were engineered to production standards. In the case of Thinking Machines, copies of the CM-2 and CM-200 machines placed by ARPA were essentially refinements of the original CM-1 design. They did not represent any significant design departure from the earlier machine. However, ARPA facilitated a number of CM-2 and CM-200 acquisitions.

Researchers and industry representatives are also concerned that ARPA's wide placement of Intel and Thinking Machines products has made it difficult to get access to alternative MPP computers. For example, since officials at government laboratories are aware that ARPA has a program to place Intel and Thinking Machines products in laboratories, they may be more inclined to take advantage of this opportunity than to independently acquire alternative machines. Even researchers who are simply looking for access to existing MPP machines have found it much easier to get access to Intel and Thinking Machines computers than other MPP machines because the ARPA-sponsored products have been widely placed in the community. Accordingly, the research community may be less able to fully explore alternative designs because a lack of ARPA support has limited the variety of MPP machines available at laboratories.

Recognizing the concerns regarding its narrow placement of MPP computers, ARPA has begun to experiment with an alternative approach. In a January 1992 solicitation, ARPA offered to include qualified HPC developers who do not already have development contracts with ARPA in its prototype placement program.[2] However, contracts stemming from this provision do not specifically guarantee sales of any particular vendor's products but merely make a vendor eligible for consideration. Although ARPA is currently negotiating contracts related to this provision with several vendors, no computer placements have yet been made.

Conclusions

It is important that mechanisms be in place to provide new, advanced computers to the research community in a rapid, streamlined fashion. ARPA has been addressing that need, but only with machines it has helped to develop. Since it appears that at least some of the ARPA placements of machines were primarily for research into the applicability of MPP rather than primarily for rapid feedback on the design of the machines, ARPA does not appear to be justified in restricting the program to only those machines that it helped to develop. A broader and richer range of experimental results could be obtained by placing a wider range of machine designs. ARPA has already begun experimenting with a more broadly based approach to placement of machines. ARPA's goal of stimulating the aggressive advance of the high-performance computing technology base would be best served by adopting such an approach for all of its MPP placements.

[2]Broad Agency Announcement 92-07, "Research in High Performance Computing Systems," Commerce Business Daily, Jan. 27, 1992.

Recommendations

We recommend that the Secretary of Defense direct the Director, ARPA, to (1) broaden the agency's computer placement program by including a wider range of computers from more vendors; (2) ensure that the agency's decisions to support high-performance computing research projects are not influenced by whether ARPA-supported computing platforms have been specified in researchers' proposals; and (3) develop, maintain, and publish a comprehensive list of qualified machines, including both ARPA-sponsored and non-ARPA-sponsored designs, that ARPA will assist sponsored researchers to acquire.

Agency Comments

ARPA officials stressed that they are already taking steps to broaden their placement of MPP machines. Specifically, they said that the alternative approach to placing computer systems that they are experimenting with represents significant progress in broadening the range of machines the agency places. As discussed in the report, we recognize that ARPA is currently negotiating contracts related to this provision with a number of vendors. However, no computer placements have yet been made. We believe that attention to this issue is still needed to ensure that a broader program ensues.

Also, ARPA officials stated that the agency has recently been involved in a joint project with NSF to place computer equipment made by a current rival to Intel and Thinking Machines. Specifically, the agency funded enhancements to two existing computer systems manufactured by Kendall Square Research Corporation, which are located at Cornell University and the Oak Ridge National Laboratory. However, in our opinion, these enhancements are not comparable to the placement of complete new systems. The totals that we report for Intel and Thinking Machines were for the purchase and installation of new machines, not upgrades or enhancements to existing installations.

ARPA's HPC Program Would Benefit From Greater Collaboration With the HPCC Research Community

Despite its widely recognized leadership role in HPC research and development, ARPA's program remains isolated from the larger HPCC research community. Little information is published about the direction of the program or the projects that are being funded, leaving potential participants uncertain about ARPA's specific technical priorities. ARPA's process for reviewing and selecting proposals is also somewhat informal and difficult for observers to follow. Since project results are not required to be reported at any predetermined time, up-to-date performance data on new MPP designs developed by ARPA have been hard to obtain. The lack of such data hinders wider adoption of MPP technology because the scientific community remains uncertain about the merits of new designs. To stimulate rapid HPC progress, ARPA will need to interact more effectively with the broader HPCC community and disseminate information about the objectives and results of its program expeditiously, including timely performance data on new MPP designs.

ARPA's Interaction With the Research Community Is Too Limited

Although ARPA funding plays a key role in the HPC community, the agency obtains little input from the community on the direction of its program and disseminates little public information about its projects and their status. HPC researchers without ARPA contracts have limited opportunity for input into the technical direction of the ARPA program. Program managers within the agency formulate HPC goals and priorities based on their knowledge of the field. According to ARPA officials, the chief mechanism for obtaining input on the overall direction of the HPC program is through discussions with ARPA-sponsored researchers at principal investigator meetings. Forums that would allow for direct input from non-ARPA sponsored researchers do not exist, and attendance at principal investigator meetings has been criticized by researchers as not representative of key research and development constituents who have significant stakes in the overall HPCC program. A more open meeting, such as an annual or semi-annual conference on the direction of the HPC program, would allow for broader community involvement and promote sharing of research experience.

ARPA has also disseminated little public documentation on the HPC projects it funds and the rationale for this funding. ARPA formerly issued annual reports for its Strategic Computing Program that provided overviews of the projects being funded, budget allocations, and results obtained to date. However, these annual reports, which were issued from 1985 through 1988, have been discontinued. ARPA officials said that they decided to discontinue the annual reports because researchers were using the reports as guidance in preparing new proposals. ARPA wanted researchers instead

to see the program as a "clean slate," open to any creative and original proposal. However, the resulting lack of public information has led to confusion among potential bidders about the specific direction of ARPA's program.

Potential bidders have also had trouble learning the outcome of ARPA's project selection process.[1] Specifically, ARPA does not do enough to notify bidders of results in a timely and orderly way. A dozen or more proposals may be selected from the results of a single broad agency announcement (BAA), and these are turned over to a variety of government contracting agencies to be separately negotiated. Because each contracting agency then proceeds at its own pace, there is no convenient, single point in time when all the contracts resulting from a given BAA are announced together.

Furthermore, ARPA does not attempt to give a consolidated public accounting of the results of each BAA. As a result, it is difficult for bidders whose proposals are rejected to learn what projects were selected in their place or why. This lack of information, in turn, reinforces perceptions among some members of the MPP community that ARPA favors certain bidders and is unwilling to work with others. Better information dissemination about BAA selection results could encourage wider participation and could allay questions about ARPA's fairness.

Project Results Are Not Published in a Uniform and Timely Fashion

ARPA does not require its researchers to report on the results of their projects in any standard format or at any predetermined time. Although researchers often publish papers and articles describing project results, there is no coordinated source for information on the status of ongoing HPC projects. There may be no way to obtain performance data on a new system supported by ARPA until perhaps a year or more after the new machine has been introduced. This results in uncertainty about the technical merits of new designs and adds risk to MPP procurement decisionmaking at government laboratories.

A case in point is the Connection Machine CM-5, manufactured by Thinking Machines Corporation. The CM-5 was a major new MPP product that attracted attention in the research community. Although its design theoretically promised significantly greater processing speed than older MPP machines, actual performance was hard to predict in the abstract. Without quantitative measurement of the machine's performance running

[1]Appendix III contains a full analysis of ARPA's HPC project selection process and discusses the results of a sample case.

real applications, researchers said they had difficulty determining the CM-5's true capabilities. For example, one NASA official involved in planning MPP research said that he had heard comments about the CM-5's performance ranging from "terrific" to "atrocious" but had seen no quantitative data to support either claim. Another noted that the lack of performance data on the machine had initially made researchers skeptical about the CM-5's capabilities. Although many researchers consider the CM-5 to be a significant advance in MPP development, quantitative performance data still remain scarce.

MPP Performance Measurement Standards Are Lacking

There are no widely accepted methods for measuring and comparing the performance of different MPP systems. Although a variety of benchmark problem sets have been developed that typify the problem domains of specific scientific and engineering disciplines, these benchmarks are often only useful within those disciplines. As a result, the performance of these machines in other domains is unknown, which makes it difficult for researchers who are interested in applying MPP to their work to gauge progress in the field.

ARPA does not support the use of universal benchmarks to evaluate and compare the performance of MPP systems, arguing that the difference among MPP designs is too great for a single suite of benchmark tests to adequately profile. ARPA officials believe that potential MPP users should instead test machines by using ARPA-sponsored libraries of software modules developed for their disciplines. ARPA argues that trying out these software modules would give prospective users a good idea of how well a given machine would perform. While this could theoretically take the place of benchmarks for some users, a mechanism is still needed to compile comparable performance results across platforms and make them readily available. Advocates of such benchmarks maintain that a standard performance metric is needed to force an objective assessment of the relative merits of different MPP systems. They argue that such an assessment would allow potential MPP users to make more intelligent investment choices and would focus MPP research and development efforts on overcoming performance shortfalls.

Even if standard performance metrics remain difficult to establish, ARPA could nevertheless improve the amount and quality of performance data that are available to the research community. Some experts have suggested that a consolidated database be established that would offer researchers easy access to information about the status and results of

ongoing HPCC projects, including performance assessments of new supercomputer designs. Establishing such a database would go far to resolve ARPA's information dissemination shortcomings. In addition to providing potential users of MPP technology with data on which to base their investment decisions, such a database would mark progress in both hardware and software toward the HPC program's goals and provide feedback to the community on research results and areas in need of emphasis or redirection.

Conclusions

ARPA's leadership role in HPC research and development carries with it the responsibility to address the needs of researchers as broadly and as fairly as possible. Unaccustomed to dealing with such a broad constituency, ARPA has focused on interacting chiefly with its own principal investigators. As a result, the agency has not devoted enough attention to disseminating information about its HPC program and its project selection process to the larger HPCC community. ARPA needs to widen its contact with this community, disseminating information more expeditiously and soliciting more input regarding the direction of its program. Establishing a database of performance and program data would go far to resolve the agency's information dissemination shortcomings.

Recommendations

We recommend that the Secretary of Defense direct the Director, ARPA, to (1) establish and maintain a public database of information about the status and results of the agency's ongoing HPC projects as well as performance data for different MPP systems, and (2) sponsor annual or semi-annual conferences on the direction of the HPC program specifically to provide a forum for broad scientific input to the agency's HPC program.

Agency Comments

ARPA officials said they are considering alternatives for improving information dissemination within the HPC program, such as establishing a public database of program information. However, they stated that they do not consider their current information dissemination activity to be lacking. In support of their contention, they referred to several public conferences on supercomputing that have been held where general information about the governmentwide HPCC program—including ARPA's component—was disseminated. They also cited published materials associated with their principal investigator meetings. However, we remain concerned that the amount and level of detail of the information provided through these mechanisms is inadequate. Information about the scope and direction of

the ARPA program was presented at the public conferences at only a very abbreviated and generalized level. Likewise, the published results of ARPA principal investigator meetings have consisted merely of collections of presentation graphics, with no explanatory material that would make them useful to individuals who were unable to attend the meetings. Furthermore, none of these mechanisms for information dissemination addresses the need to obtain broad scientific input to the agency's HPC program. Nor does any mechanism offer a way to ensure that detailed performance results from critical new technology projects are disseminated expeditiously.

ARPA Must Emphasize Software to Meet Its Goals

ARPA must address critical software shortcomings quickly in order to meet its HPC program goal of making MPP useful in addressing a variety of complex real-world problems. Although hardware systems capable of scaling up to 1 trillion operations per second appear to be achievable by 1996, it is much less likely that the system software required for users to harness this power will be available at that time. System software development has not kept pace with hardware development, leaving current systems difficult to use. User groups studying this problem have identified MPP programming languages, advanced compiler systems, and program development tools as specific areas that are critical to making MPP systems more useful. Unless ARPA shifts the balance of its program to emphasize these software areas, it appears unlikely that the capabilities of advanced MPP supercomputers will be fully exploitable by 1996.

Experts Have Emphasized the Need for Greater Investment in System Software

There is widespread concern among university and industry experts that software shortcomings are holding back further advances in MPP. These experts argue that MPP will not gain widespread acceptance unless rapid progress is made in developing system software. Such progress can only be achieved by investing more heavily in all phases of HPC software research and development.

Three separate conferences convened in 1992 to identify critical directions in research and development for HPC have emphasized the critical role of system software in ensuring the success of the HPCC program. The Purdue Workshop on Grand Challenges in Computer Architecture for the Support of High Performance Computing, sponsored by the National Science Foundation, stressed that a substantial effort must be devoted to advancing MPP system software.

Similarly, the Pasadena Workshop on System Software and Tools for High Performance Computing Environments, sponsored by a number of federal agencies and attended by over a hundred experts from industry, academia and government, concluded that a substantial investment of resources and time in the research and development of system software will be required to make MPP hardware systems useful. Scientists participating in an Industry Advisory Board forum, convened by International Business Machines and the Association for Computing Machinery, also concluded that the lack of system software is the most serious obstacle confronting MPP users.

Primitive System Software Limits MPP's Usefulness

For most potential MPP users, the high cost of developing software for their specific applications still outweighs the benefits of greatly improved processing speed associated with MPP. Application software development is difficult because current MPP systems are equipped with only rather primitive system software, thus requiring extra effort on the part of application programmers. For example, MPP programmers usually must be highly versed in the particular architectural features of their machines, such as details about how the memory is configured. In many cases, they also must specify exactly how information controlled by different processors should be exchanged. Furthermore, once they have written their programs, few tools are available to assist programmers in examining how the instructions are divided among and executed by the processors so that errors can be isolated and performance improved. As a result, only the most highly skilled potential users can effectively exploit the performance capabilities of most MPP systems. In addition to the immediate negative impact on scientific researchers, the lack of system software also discourages vendors from developing commercial MPP application software, because such products may only work on a few specific hardware platforms that may quickly become obsolete.

There is general consensus in the HPC community that several system software technologies must be researched and developed soon in order to ease the burden on programmers and accelerate conversion to MPP. These technologies include: (1) parallel programming languages and compilers and (2) software development tools.

Better Parallel Programming Languages and Compilers Are Needed

Programmers are often unable to take full advantage of the capabilities of MPP machines because the programming languages and compilers available today do not have adequate functionality. Although development of software tends to lag behind hardware development for all types of computer systems, certain factors have made development of parallel programming languages and compilers especially difficult. Since MPP designs are widely divergent, each with its unique organization of multiple processors and memory hierarchies, it is especially difficult to develop system software that can efficiently interact with each design. As a result, most current MPP computers have relatively weak system software that does not perform all the functions that it should. For example, MPP programmers must explicitly synchronize and schedule concurrent program activities and resources, a complex and arcane task that consumes a great deal of programming effort and is often beyond the experience of most application programmers.

Existing parallel programming languages and compilers generally cannot carry out important functions efficiently and effectively or differ widely in the approaches they take. Examples of such functions include (1) how a user controls parallel input and output, (2) the way in which processors that are working in parallel on a problem communicate among themselves, and (3) how data are to be arranged in memory.

Computer scientists acknowledge that much work needs to be done to develop languages and compilers that would automatically address all these unique MPP functions in a highly efficient and effective way, and they agree that this is an area that should dominate future research and development efforts.

Better Software Development Tools Are Needed

Parallel program debuggers are foremost among the tools that are needed to enable users to develop MPP application software efficiently. Parallel programs are more difficult to debug than traditional programs because of the many processors involved, each of which may independently encounter programming errors while other processors are executing software instructions correctly. The problem is complicated by the fact that erroneous results from one processor are likely to be fed into computations performed by other processors. The improper timing of activities shared among several processors can cause errors that are hard to detect because they do not occur in the same way every time the program is run. Debugging facilities capable of diagnosing these timing-dependent errors and relating them back to the programmer's code have not yet been developed. The lack of such tools severely hinders MPP users' efforts to design and develop parallel applications.

Even if a parallel program is free of errors, it may not run as quickly or efficiently as the programmer expects. The problem may be due to the user's inexperience in allocating parallel hardware resources to computing tasks. Tools that enable users to interact with and understand the behavior of their programs are essential for producing programs that make the most efficient use of MPP hardware resources. Preferably, these tools would provide a graphic analysis and explanation so that the programmer could visualize the problem and more readily correct it.

ARPA's Approach Focuses on Early Phases of Software Development

ARPA officials recognize that software issues must be resolved in order to reach their goal of enabling the application of MPP to the grand challenges by 1996. They stress that the objective of the ARPA HPC program is to develop MPP systems, including hardware and software. The ARPA program includes a system software component intended to enhance the practical usefulness of MPP systems. Compilers, programming languages, and software development tools are all being addressed to some degree within this program.

Nevertheless, users with whom we spoke were almost unanimous in criticizing ARPA's support for parallel software research and development as inadequate. Although ARPA is involved in software research and development, its commitment to making the same kind of advances in software as it has made in hardware is not apparent to researchers. This may be in part because ARPA's strategy for software research and development is to invest most in early work that explores new software concepts and involves experimentation with laboratory models. ARPA's policy is not to spend heavily on the later, advanced stages of software development, which concentrate on making reliable products for widespread use.

Experts working on MPP software say that ARPA's strategy is inadequate for the current state of the MPP industry. They maintain that ARPA needs to emphasize software development by continuing to fund the most promising of the research prototypes through the advanced development stage. Readying more software products for evaluation by end users is seen as the best way to accelerate the community's acceptance of MPP. Furthermore, commercial software developers, who heretofore have been reluctant to invest heavily in the development of MPP software, would likely be encouraged to produce more refined commercial software products. This is because ARPA support for any given technology is often seen by commercial developers as leading to greater acceptance and more widespread utilization by the user community.

Conclusions

Current MPP software shortcomings are significant. Experts in the field agree that these shortcomings are a barrier to effective utilization of MPP within the scientific community. Specific software areas in need of intensive near-term research and development have been identified by expert panels. However ARPA has not provided strong leadership nor has it made a clear commitment to accelerating system software research and development efforts. If ARPA is to meet its goal of enabling the effective

application of MPP to grand challenge problems by 1996, it will need to shift its emphasis to software.

Recommendations

We recommend that the Secretary of Defense direct the Director, ARPA, to (1) emphasize and support the research and development of MPP system software as a major element of the agency's HPC program, and (2) ensure that a significant body of software development work is underway to aggressively advance the technology base, especially in MPP programming languages, effective compiler systems, and debugging and performance analysis tools. In order to achieve broad MPP use by 1996, ARPA must fund promising software prototypes to an advanced stage of development as it does for hardware, rather than limiting itself to supporting only the earlier stages of research and development.

Agency Comments

ARPA officials disagreed with our statement that ARPA is not sufficiently addressing MPP system software. They stated that they have made substantial investments in software development, including providing funding to most prominent software researchers in the field. We do not dispute that ARPA has funded a substantial number of software development projects. However, it was clear from our review that the level of system software investment that has been made heretofore has not succeeded in creating stable, reliable computing platforms that scientists can use to address grand challenge problems. Researchers we contacted were in broad agreement that progress in this area will need to accelerate if ARPA is to reach its objective of making MPP systems useful to researchers by 1996. Accordingly, greater emphasis on developing system software is still needed.

Objectives, Scope, and Methodology

In its report accompanying the National Defense Authorization Act for Fiscal Year 1993, the House Committee on Armed Services requested that we assess the investment strategy of the ARPA HPC program. During subsequent meetings with committee staff, our specific objectives were established. These objectives were to assess (1) ARPA's distribution of advanced computers to research sites, (2) ARPA's interaction with the research community, and (3) the balance between software and hardware investments in the ARPA program.

To meet these objectives, we obtained and reviewed official program documentation and discussed these issues with government, private industry, and academic officials from a wide range of organizations.

Specifically, with regard to ARPA's role within the federal HPCC program, we interviewed officials at:

- Office of Science and Technology Policy, Executive Office of the President, Washington, D.C.;
- National Coordination Office for HPCC, National Library of Medicine, Bethesda, Maryland;
- National Science Foundation, Directorate for Computer and Information Science and Engineering, Washington, D.C.;
- Department of Energy, Office of Energy Research, Gaithersburg, Maryland; and
- Department of Commerce, International Trade Administration, Washington, D.C.

Regarding the specifics of the ARPA program, we analyzed summary records of computer procurements and ARPA's project selection process and interviewed officials at:

- ARPA, Computing Systems Technology Office, Arlington, Virginia;
- Office of the Director of Defense Research and Engineering, Washington, D.C.;
- National Research Council, Computer Science and Telecommunications Board, Washington, D.C.; and
- National Institute of Standards and Technology, Computer Systems Laboratory, Gaithersburg, Maryland.

We also interviewed scientific researchers and officials at government laboratories, including:

- National Security Agency, Ft. Meade, Maryland;
- Army High Performance Computing Research Center, Minneapolis, Minnesota;
- David Taylor Research Center, Carderock, Maryland;
- Naval Research Laboratory, Washington, D.C.;
- Naval Command, Control and Ocean Surveillance Center, San Diego, California;
- San Diego Supercomputer Center, San Diego, California;
- Lawrence Livermore National Laboratory, Livermore, California;
- Los Alamos National Laboratory, Los Alamos, New Mexico;
- Sandia National Laboratory, Albuquerque, New Mexico;
- NASA Ames Research Center, Moffett Field, California; and
- NASA Goddard Space Flight Center, Greenbelt, Maryland.

We interviewed representatives from the academic community at:

- National Science Foundation, Washington, D.C.;
- Stanford University, Palo Alto, California;
- University of Maryland, College Park, Maryland;
- Massachusetts Institute of Technology, Cambridge, Massachusetts; and
- Minnesota Supercomputer Center, Minneapolis, Minnesota.

We interviewed MPP industry officials representing:

- Cray Research, Inc.;
- Intel Corporation, Supercomputer Systems Division;
- Kendall Square Research Corporation;
- MasPar Computer Corporation;
- nCUBE Corporation; and
- Thinking Machines Corporation.

We also reviewed reports and documents related to HPC prepared by various advisory groups and committees and attended a discussion of HPC organized by the American Electronics Association in Washington, D.C.

We conducted our review from June 1992 to February 1993, in accordance with generally accepted government auditing standards.

ARPA's Placement of Intel and Thinking Machines Products

Table II.1: Intel Touchstone Machines Placed by ARPA

Date	Site	Machine Type	Acquired	Number of Nodes	Cost	Source of Funding
10-30-89	NASA	gamma	new	128	$1,534,100	NASA
06-07-90	NSA	gamma	new	16	$445,949	NSA
07-17-90	NIH	gamma	new	16	$501,117	NIH
07-20-90	Wright-Patterson Air Force Base	gamma	new	three 8 one 16	$620,020 $518,938	Air Force Army
08-03-90	NASA	gamma	new	two 8	$399,834	NASA
08-20-90	Naval Ocean Systems Center	gamma	new	16	$380,000	Navy
08-20-90	NSA	gamma	upgrade	16	$333,960	NSA
12-05-90	Naval Coastal Systems Center	gamma	new	one 32 one 16	$840,000	Navy
12-10-90	Stanford University	gamma	new	32	$604,300	ARPA
03-21-91	Naval Research Laboratory, Phillips Laboratory, Wright-Patterson Air Force Base, Ballistics Research Laboratory	gamma	new	seven 8	$1,333,983	ARPA
03-22-91	San Diego Supercomputer Center, Cornell Theory Center	gamma	new	two 32	$1,016,730	ARPA
04-01-91	ARPA	gamma	new	three 8 one 32	$1,036,500	ARPA
04-03-91	Naval Ocean Systems Center	gamma	upgrade	16	$83,415 $145,000	ARPA Navy
05-30-91	NSA	gamma	upgrade	32	$637,849	NSA
05-31-91	NIH	gamma	upgrade	64	$952,688	NIH
06-04-91	Naval Research Laboratory	gamma	upgrade	two 8	$131,560	ARPA
06-06-91	ARPA	gamma	new	two 8	$291,895	ARPA
08-16-91	Naval Coastal Systems Center	gamma	upgrade	16	$231,165	Navy
08-16-91	Naval Research Laboratory	gamma	upgrade	8	$137,595	Navy
08-16-91	Sandia National Laboratories	gamma	upgrade	24	$496,096	DOE
09-04-91	Phillips Laboratory	gamma	upgrade	8	$11,325	Air Force
09-11-91	NASA	gamma	upgrade	16	$285,655	NASA
12-11-91	NIH	gamma	upgrade	48	$936,097	NIH
12-12-91	DOE	gamma	new	64	$995,000	DOE
02-28-92	Naval Coastal Systems Center	gamma	upgrade	combined two 32 to one 64	$151,860	Navy
04-01-92	NIH	gamma	upgrade	64	$459,384	NIH
05-13-92	Naval Research Laboratory	gamma	upgrade	8	$121,000	Navy
05-20-92	NSA	sigma	new	60	$1,111,500	NSA
09-01-92	NASA	sigma	new	28	$3,367,500	NASA
09-11-92	ARPA	sigma	new	60	$1,425,936	ARPA

Note: 32 Touchstone prototypes have been placed by ARPA, not including upgrades.

Table II.2: Intel iWarp Machines Placed by ARPA

Date	Site	Machine Type	Acquired	Number of Nodes	Cost	Source of Funding
4-17-90	Naval Ocean Systems Center	iWarp	new	two 64	$260,000 $540,000	ARPA Navy
02-07-91	Naval Undersea Warfare Center	iWarp	new	32	$241,380	ARPA
03-26-91	Naval Undersea Warfare Center	iWarp	new	64	$481,350	Navy
04-29-91	University of Maryland	iWarp	new	two 32 one 16	$471,660 $143,550	ARPA
07-91	Carnegie Mellon University	iWarp	new	three 64	$1,091,250 $615,110	ARPA Strategic Defense Initiative Organization
07-03-91	NSA	iWarp	new	64	$420,300	NSA
03-03-91	Naval Air Warfare Center	iWarp	new	64	$50,240 $400,000	ARPA Navy
03-03-92	Naval Undersea Warfare Center	iWarp	upgrade	32	$221,700	Navy

Note: Twelve iWarp prototypes have been placed by ARPA, not including upgrades. In addition, ARPA funded the placement of 16 eight-node development systems at a total cost of $2,007,660, for experimentation and code development by government researchers and contractors.

Table II.3: Thinking Machines Corp. Connection Machines Placed by ARPA

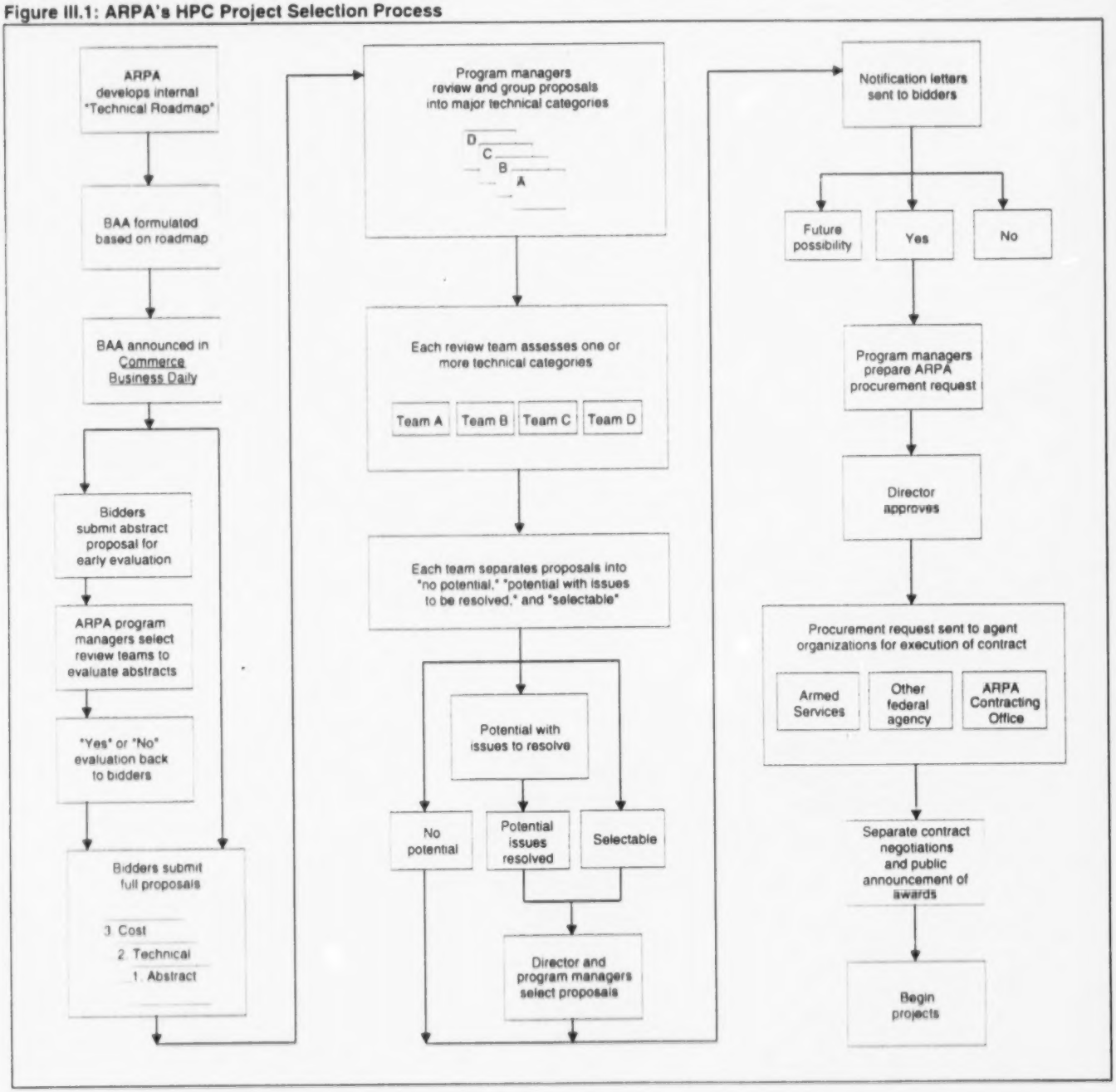
Date	Site	Machine Type	Acquired	Number of Nodes	Cost	Source of Funding
	Network Server	CM1	new	16,000		
	Naval Research Laboratory	CM1	new	16,000		
08-29-86	Science Applications International Corporation	CM1	new	16,000	$2,820,000	ARPA
	Syracuse University	CM1	new	16,000		
	Lockheed Corporation	CM2	new	two 16,000		
	Syracuse University	CM2	new	16,000		
06-24-87	SRI International	CM2	new	8,000	$3,350,000	ARPA
	The Analytical Sciences Corporation	CM2	new	8,000		
	University of Maryland	CM2	new	16,000		
05-16-88	Naval Research Laboratory	CM2	new	16,000	$2,483,113 2,072,202	ARPA Strategic Defense Initiative Organization
	University of Southern California	CM2	new	16,000		
	NASA	CM2	new	16,000		
06-88	National Center for Atmospheric Research	CM2	new	8,000	$625,000	ARPA
08-88	Sandia National Laboratories	CM2	new	16,000	$1,109,005	ARPA
11-89	Naval Ocean Systems Center	CM2	new	8,000	$610,000	Navy
	Advanced Decision Systems	CM2	new	8,000		
11-14-89	Engineering Topographic Laboratory	CM2	upgrade	8,000	$1,100,000	ARPA
	Science Applications International Corporation	CM2	upgrade	8,000		
	Naval Research Laboratory	CM2	upgrade	8,000		
12-89	Naval Undersea Warfare Center	CM2	new	8,000	$835,000	Navy
	Harvard University	CM2	new	4,000		
12-18-89	SRI International	CM2	new	4,000	$1,827,000	ARPA
	Naval Ocean Systems Center	CM2	new	8,000		
05-16-91	NASA	CM2	upgrade	16,000	$1,960,338	NASA
12-88	Naval Research Laboratory	CM2	various upgrades	16,000	$850,443	Navy
11-92	Naval Research Laboratory	CM5	new	128	$100,000 $4,833,000	ARPA Navy
12-92	ARPA	CM5	new	32	$989,724	ARPA
01-93	NASA	CM5	new	128	$4,636,740	NASA

Note: Twenty-four Connection Machine prototypes of various sizes have been placed by ARPA, not including upgrades.

ARPA's HPC Project Selection Process

Once ARPA has identified the specific technical areas that it intends to fund, ARPA formulates one or more broad agency Announcements (BAA), which state its intention of investing in a particular field and describe in general terms its near-term research goals, the criteria for proposal selection, the method of evaluation, the deadline for submitting proposals, and instructions on how proposals should be prepared and submitted. ARPA then releases the BAA to the research community by publishing it in the Commerce Business Daily. ARPA BAAs may be jointly announced with other federal agencies.

Figure III.1: ARPA's HPC Project Selection Process


ARPA develops internal "Technical Roadmap"
BAA formulated based on roadmap
BAA announced in Commerce Business Daily
Bidders submit abstract proposal for early evaluation
ARPA program managers select review teams to evaluate abstracts
"Yes" or "No" evaluation back to bidders
Bidders submit full proposals
3. Cost
2. Technical
1. Abstract
Program managers review and group proposals into major technical categories
D
C
B
A
Each review team assesses one or more technical categories
Team A
Team B
Team C
Team D
Each team separates proposals into "no potential," "potential with issues to be resolved," and "selectable"
Potential with issues to resolve
No potential
Potential issues resolved
Selectable
Director and program managers select proposals
Notification letters sent to bidders
Future possibility
Yes
No
Program managers prepare ARPA procurement request
Director approves
Procurement request sent to agent organizations for execution of contract
Armed Services
Other federal agency
ARPA Contracting Office
Separate contract negotiations and public announcement of awards
Begin projects

Prior to developing and submitting a full proposal, responding organizations may submit an abstract proposal to get an early indication of whether their proposals are of interest to ARPA. Teams of officials, largely from ARPA but also including technical experts from other agencies, then meet to review the abstracts. Feedback on the abstract proposals is in the form of a simple "yes" or "no," although a "no" response does not preclude later submission of a full proposal.

After all proposals have been submitted, they are reviewed and grouped into major technical categories, which teams of officials evaluate based on ARPA's technical priorities and the proposals' estimated costs. The reviewers then sort the proposals into three categories: selectable, potentially selectable with technical issues to resolve, and no potential to negotiate. ARPA staff then attempt to resolve the unresolved technical issues by contacting the bidders and obtaining clarification about their bids. Once this process is complete, the HPC Program Director and the Director's staff select proposals for funding.

Following the selection of the proposals, notification letters are sent to the bidders stating that either their proposal (1) has been selected for negotiation, (2) has potential for future ARPA funding, or (3) was not selected for award.

Program managers then prepare an ARPA Procurement Request for selected proposals. This request is a one- or two-page description of the project and how it relates to ARPA's program. It also describes how ARPA plans to fund the project. The procurement request is then given to the HPC Program Director for approval.

After this entire procedure is complete, ARPA sends the approved proposal either to ARPA's Contract Management Office or to another federal agency, which then negotiates the final contract with the bidder.

Under any given BAA, a variety of contracts ranging in scope and size may be signed, all at different times, through a variety of contracting agencies. About two-thirds of the contracts ARPA awards are drawn up by other agencies. Only a few of the larger, more complex contracts are directly managed by ARPA's Contract Management Office, which negotiates the contracts directly with bidders. ARPA's selection of any given project is only made public when a contract is signed. According to ARPA officials, 95 percent of awards will be made within 1 year of the deadline for proposal submissions.

A Case Study: Synopsis #57

Synopsis #57, one of the earliest BAAs issued within the Strategic Computing Program, was published in the Commerce Business Daily on April 1, 1987, by ARPA and the Space and Naval Warfare Systems Command to solicit proposals for research in the area of parallel computing systems. The announcement was open for 3 months after publication. ARPA received 38 proposals in response to Synopsis #57. These proposals were sorted into five different technical categories: (1) General Purpose Processors, (2) Very Large-scale Integration and Packaging, (3) Processor Studies and Support, (4) Signal and Systolic Processors, and (5) proposals spanning more than one of the other categories. Proposals in the first four categories were evaluated by teams that weighed all the submissions in each category against each other. Proposals in the fifth category were assessed individually by ARPA officials.

ARPA awarded 15 contracts: two in the General Purpose Processor category, three in the Very Large-scale Integration and Packaging category, three in the Processor Studies and Support category, and none in the Signal and Systolic Processor category. In addition, seven proposals from the fifth category were selected. These contracts were awarded between June 1987 and September 1988 and had a combined cost of about $55.7 million. Table III.1 lists the awardees and the total amounts of the contracts.

Table III.1: Proposals Awarded Under Synopsis #57

Awarded Bidder	Date	Contract Total
California Institute of Technology	June 1987	$4,120,650
Stanford University	August 1987	$6,443,579
Massachusetts Institute of Technology	September 1987	$6,377,000
Stanford University	December 1987	$1,493,277
Syracuse University	January 1988	$4,932,360
Intel Corporation	March 1988	$9,300,000
Microelectronics and Computer Technology Corporation	April 1988	$7,623,544
University of Southern California	May 1988	$5,151,349
Princeton University	July 1988	$1,496,895
Massachusetts Institute of Technology	July 1988 July 1988 (two contracts)	$3,653,083
University of California, Santa Barbara	July 1988	$1,329,850
ESL, Incorporated	August 1988	$1,685,650
Stanford University	August 1988	$1,853,490
Syracuse University	September 1988	$279,994

Note: Intel's Touchstone project, which was one of the seven individually evaluated proposals, was awarded under Synopsis #57. It was a 36-month effort costing $9.3 million, the largest contract ARPA awarded under this BAA. ARPA awarded the last contract under this announcement in September 1988. All except one of these contracts have been completed.

Major Contributors to This Report

Information Management and Technology Division, Washington, D.C.

Franklin W. Deffer, Project Director
John A. de Ferrari, Project Manager
Dr. Rona B. Stillman, Chief Scientist
Elizabeth L. Johnston, Computer Scientist
Nancy M. Kamita, Computer Scientist
Lynne L. Goldfarb, Publishing Adviser

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are $2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20884-6015

or visit:

Room 1000
700 4th St. NW (corner of 4th and G Sts. NW)
U.S. General Accounting Office
Washington, DC

Orders may also be placed by calling (202) 512-6000 or by using fax number (301) 258-4066.

PRINTED ON RECYCLED PAPER

United States
General Accounting Office
Washington, D.C. 20548

Official Business
Penalty for Private Use $300

First-Class Mail
Postage & Fees Paid
GAO
Permit No. G100